

# Network Measurement

US ATLAS Tier2/Tier3 Workshop

Aug 19-20, 2009

Aaron Brown

[aaron@internet2.edu](mailto:aaron@internet2.edu)

# Tier-2 On-Demand Testing Services

- Tier-2s are in the process of deploying software and hardware that will allow Tier-3s to perform on-demand tests to check their performance
  - Throughput
  - Latency/Jitter
- The client software for using these testing services is available via VDT 2.0.0

# Tier-2 On-Demand Testing Services

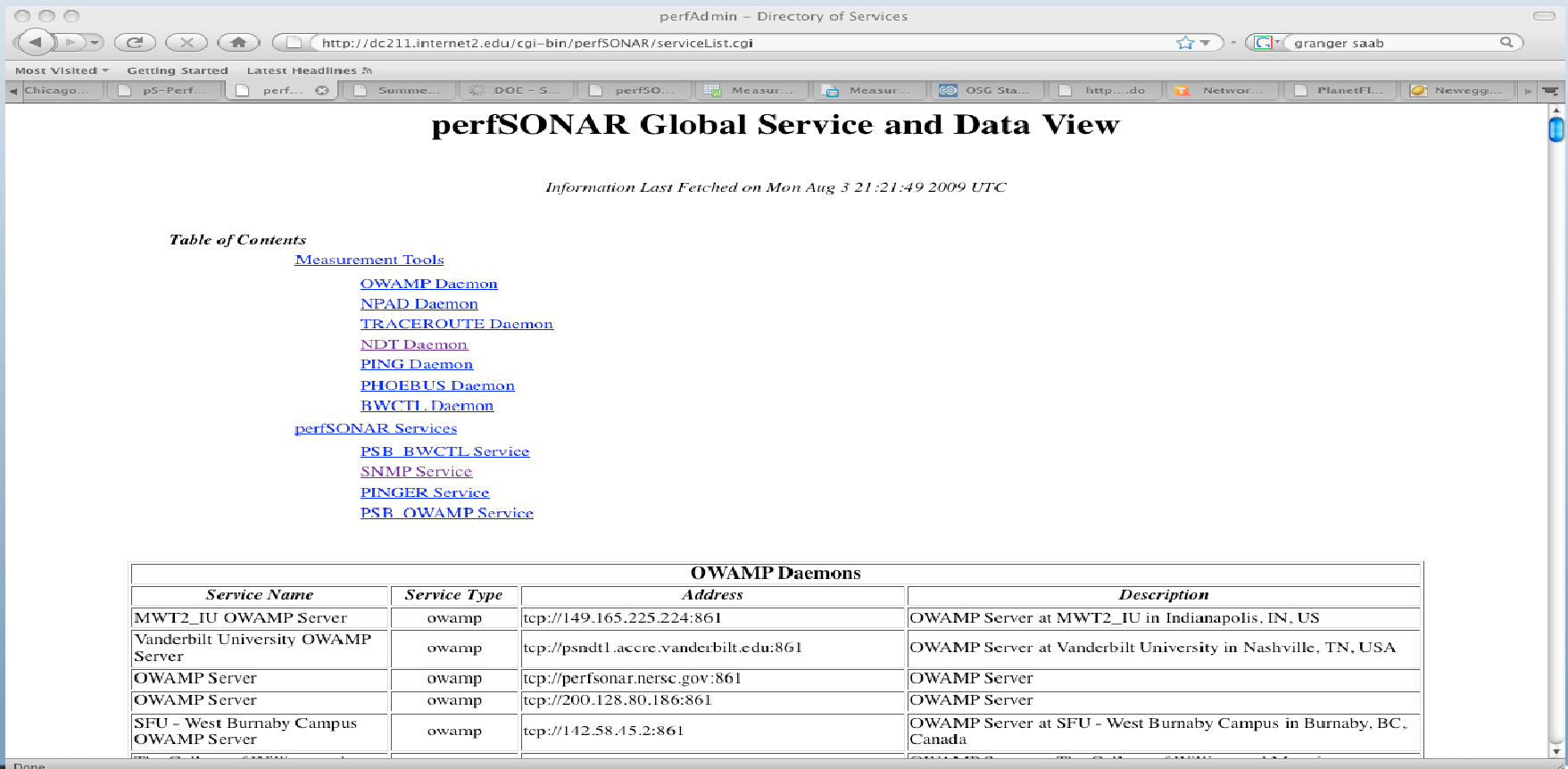
- Throughput Testing (bwctl)
  - Wraps 'iperf' bandwidth tester
  - Schedules tests so they won't interfere with other tests
  - Allows limiting the tests users can run
  - <http://www.internet2.edu/performance/bwctl/>
- One-way Delay Testing (owamp)
  - One-way "ping" command
  - Low bandwidth; tests don't interfere with each other
  - <http://www.internet2.edu/performance/owamp/>

# Tier-2 On-Demand Testing Services

- Network Diagnostic Tool (NDT), Network Path and Application Diagnostic (NPAD)
  - Measure performance to users desktop
    - Simple to use and understand Java applet
  - Identify real problems for real users
    - Network infrastructure is the problem
    - Host tuning issues are the problem
  - NPAD
    - <http://www.psc.edu/networking/projects/pathdiag/>
  - NDT
    - <http://www.internet2.edu/performance/ndt/>

# Finding deployed services to test with

- <http://www.perfsonar.net/activeServices>



perfAdmin - Directory of Services

http://dc211.internet2.edu/cgi-bin/perfSONAR/serviceList.cgi

granger saab

Most Visited Getting Started Latest Headlines 5

Chicago... p5-Perf... perf... Summe... DOE - S... perfSO... Measur... Measur... OSG Sta... http...do Networ... PlanetFl... Newegg...

## perfSONAR Global Service and Data View

Information Last Fetched on Mon Aug 3 21:21:49 2009 UTC

**Table of Contents**

- [Measurement Tools](#)
  - [OWAMP Daemon](#)
  - [NPAD Daemon](#)
  - [TRACEROUTE Daemon](#)
  - [NDT Daemon](#)
  - [PING Daemon](#)
  - [PHOEBUS Daemon](#)
  - [BWCTL Daemon](#)
- [perfSONAR Services](#)
  - [PSB\\_BWCTL\\_Service](#)
  - [SNMP Service](#)
  - [PINGER Service](#)
  - [PSB\\_OWAMP\\_Service](#)

OWAMP Daemons			
Service Name	Service Type	Address	Description
MWT2_IU OWAMP Server	owamp	tcp://149.165.225.224:861	OWAMP Server at MWT2_IU in Indianapolis, IN, US
Vanderbilt University OWAMP Server	owamp	tcp://psndt1.accre.vanderbilt.edu:861	OWAMP Server at Vanderbilt University in Nashville, TN, USA
OWAMP Server	owamp	tcp://perfsonar.nersc.gov:861	OWAMP Server
OWAMP Server	owamp	tcp://200.128.80.186:861	OWAMP Server
SFU - West Burnaby Campus OWAMP Server	owamp	tcp://142.58.45.2:861	OWAMP Server at SFU - West Burnaby Campus in Burnaby, BC, Canada

# Throughput Testing (bwctl)

```
# bwctl -i 2 -t 20 -c bwctl.losa.net.internet2.edu
```

```
# bwctl -i 2 -t 20 -s bwctl.newy.net.internet2.edu
```

- -i 2 = report intermediate results every 2 seconds
- -t 20 = run test for 20 seconds
- -s name = remote end will send data to you
- -c name = you will send data to the remote host

# Throughput Testing (bwctl)

```
rcarlson@triton:~/ndt — ssh — 116x33
-bash-3.2$ bwctl -i2 -t20 -s lhcmon.bnl.gov
bwctl: Unable to contact a local bwctld: Spawning local tool controller
bwctl: NuttcpAvailable(): We were unable to verify that nuttcp is working. Likely you do not have it installed. exit
  status: 1: output: exec(nuttcp): No such file or directory
bwctl: Couldn't initialize tool 'nuttcp'. Disabling it.
bwctl: Using tool: iperf
bwctl: 28 seconds until test results available

RECEIVER START
bwctl: exec_line: iperf -B 207.75.164.104 -s -f a -m -p 5001 -t 20 -i 2
bwctl: start_tool: 3458324284.538605
-----
Server listening on TCP port 5001
Binding to local address 207.75.164.104
TCP window size: 85.3 KByte (default)
-----
[ 12] local 207.75.164.104 port 5001 connected with 192.12.15.23 port 5001
[ 12]  0.0- 2.0 sec  19.5 MBytes  81.8 Mb/s
[ 12]  2.0- 4.0 sec  22.4 MBytes  94.0 Mb/s
[ 12]  4.0- 6.0 sec  22.4 MBytes  94.0 Mb/s
[ 12]  6.0- 8.0 sec  22.4 MBytes  94.0 Mb/s
[ 12]  8.0-10.0 sec  21.4 MBytes  89.7 Mb/s
[ 12] 10.0-12.0 sec  19.0 MBytes  79.8 Mb/s
[ 12] 12.0-14.0 sec  21.6 MBytes  90.6 Mb/s
[ 12] 14.0-16.0 sec  22.4 MBytes  94.1 Mb/s
[ 12] 16.0-18.0 sec  19.6 MBytes  82.4 Mb/s
[ 12] 18.0-20.0 sec  20.2 MBytes  84.6 Mb/s
[ 12]  0.0-20.1 sec   213 MBytes  88.5 Mb/s
[ 12] MSS size 1448 bytes (MTU 1500 bytes, ethernet)
bwctl: stop_exec: 3458324308.780934

RECEIVER END
-bash-3.2$ ]
```

# One-Way Ping (owping)

```
# owping owamp.losa.net.internet2.edu
```

- Output
  - Separate statistics for both directions
  - Number of packets sent and lost
  - One-way delay statistics min/median/max
  - Number of IP hops in path
  - Number of packets that arrive out-of-order



# One-Way Ping (owping)

```
rcarlson@nms-rlat:~ — ssh — 116x34
[rcarlson@nms-rlat ~]$ ping -c3 owamp.losa.net.internet2.edu
PING eth-1.nms-rlat.losa.net.internet2.edu (64.57.17.162) 56(84) bytes of data.
64 bytes from nms-rlat.losa.net.internet2.edu (64.57.17.162): icmp_seq=1 ttl=60 time=56.6 ms
64 bytes from nms-rlat.losa.net.internet2.edu (64.57.17.162): icmp_seq=2 ttl=60 time=56.7 ms
64 bytes from nms-rlat.losa.net.internet2.edu (64.57.17.162): icmp_seq=3 ttl=60 time=56.6 ms

--- eth-1.nms-rlat.losa.net.internet2.edu ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 2005ms
rtt min/avg/max/mdev = 56.692/56.697/56.705/0.275 ms
[rcarlson@nms-rlat ~]$ owping owamp.losa.net.internet2.edu
Approximately 13.2 seconds until results available

--- owping statistics from [2001:468:2:12::17:34]:35621 to [owamp.losa.net.internet2.edu]:55664 ---
SID: 00170162ce21d9957fb85e80c7f6958a
first: 2009-08-03T21:48:06.786
last: 2009-08-03T21:48:15.918
100 sent, 0 lost (0.000%), 0 duplicates
one-way delay min/median/max = 28.1/28.2/28.3 ms, (err=3.13 ms)
one-way jitter = 0.1 ms (P95-P50)
TTL not reported
no reordering

--- owping statistics from [owamp.losa.net.internet2.edu]:34879 to [2001:468:2:12::17:34]:35622 ---
SID: 00170034ce21d9959d6ffc1ld84cfeac
first: 2009-08-03T21:48:06.907
last: 2009-08-03T21:48:17.704
100 sent, 0 lost (0.000%), 0 duplicates
one-way delay min/median/max = 28.5/28.6/28.6 ms, (err=3.13 ms)
one-way jitter = 0 ms (P95-P50)
TTL not reported
no reordering

[rcarlson@nms-rlat ~]$
```

# NDT Client Test

File Edit View Go Bookmarks Tools Help

http://207.75.164.80:7123/

Getting Started Latest Headlines

Located at Seattle - WA; 1000 Mbps (Gigabit Ethernet) network connection

This java applet was developed to test the reliability and operational status of your desktop computer and network connection. It does this by sending data between your computer and this remote NDT server. These tests will determine:

- The slowest link in the end-to-end path (Dial-up modem to 10 Gbps Ethernet/OC-192)
- The Ethernet duplex setting (full or half);
- If congestion is limiting end-to-end throughput.

It can also identify 2 serious error conditions:

- Duplex Mismatch
- Excessive packet loss due to faulty cables.

A test takes about 20 seconds. Click on "start" to begin.

```
TCP/Web100 Network Diagnostic Tool v5.3.4e
click START to begin
Checking for Middleboxes . . . . . Done
running 10s outbound test (client to server) . . . . . 360.76Kb/s
running 10s inbound test (server to client) . . . . . 20.53Mb/s
Warning! Client time-out while reading data, possible duplex mismatch exists
The slowest link in the end-to-end path is a 100 Mbps Full duplex Fast Ethernet subnet
Alarm: Duplex Mismatch condition detected Switch=Full and Host=half

click START to re-test
```

START Statistics More Details... Report Problem

Tcpcb100 done

# NPAD Client Test

Test Results

http://speedtest2.pnl.gov:8200/ServerData/nms-rexp.seat.net.internet2.edu%3A2009-08-03-22%3A34%3A23.html

granger saab

Most Visited Getting Started Latest Headlines

Chicago... Chicago... p5-Perf... Test ... Summe... DOE - S... perFSO... Measur... Measur... OSG Sta... http...do Networ... PlanetFI... Newegg...

## Test conditions

Tester: (none) (192.101.102.24) [?]  
Target: (none) (64.57.19.4) [?]  
Logfile base name: nms-rexp.seat.net.internet2.edu:2009-08-03-22:34:23 [?]  
This report is based on a 500 Mb/s target application data rate [?]  
This report is based on a 76 ms Round-Trip-Time (RTT) to the target application [?]  
The Round Trip Time for this path section is 6.000000 ms.  
The Maximum Segment Size for this path section is 8948 Bytes. [?]

### Target host TCP configuration test: Pass! [?]

TCP negotiated appropriate options: WSCALE=10, SACKOK, and Timestamps. [?]  
The target passed all tests! See tester caveats: [?]

### Path measurements [?]

The path to the tester is too long for accurate measurements.  
> **Test a shorter path section or reduce the target data rate and/or RTT.** [?]

### Data rate test: Pass! [?]

Pass data rate check: maximum data rate was 989.456662 Mb/s [?]

### Loss rate test: Fail! [?]

Fail: loss event rate: 0.000525% (190639 packets between loss events). [?]  
Diagnosis: there is too much background (non-congested) packet loss. [?]  
The events averaged 11.700000 losses each, for a total loss rate of 0.006137%. [?]  
FYI: To get 500 Mb/s with a 8948 byte MSS on a 76 ms path the total end-to-end loss budget is 0.000174% (575094 packets between losses). [?]  
> **Locate the excess packet loss in this section of the path.** [?]

### Suggestions for alternate tests

FYI: This path may pass with a less strenuous application: [?]  
Try rate=500 Mb/s, rtt=43 ms  
Try rate=989 Mb/s, rtt=22 ms  
Or if you can raise the MTU: [?]  
Try rate=500 Mb/s, rtt=44 ms, mtu=9000 bytes  
Try rate=989 Mb/s, rtt=22 ms, mtu=9000 bytes

### Network buffering test: Warning! [?]

This test did not complete due to other problems with the path, target or tester.  
> **Correct other problems first, and then rerun this test.** [?]  
Estimated queue size is at least: Pkts: 9 Bytes: 80532  
This is probably an underestimate of the actual queue size. [?]  
This corresponds to a 0.577737 ms drain time. [?]  
To get 500 Mb/s with on a 76 ms path, you need 4750000 bytes of buffer space. [?]  
> **Localize all path problems by testing progressively smaller sections of the full path.** [?]

### Tester validation: Warning! [?]

The tester has a bottleneck. [?]  
> **Please see the instructions on tester flaws.**  
Tester version: Sld: pathdiag.py.v 1.42 2008/06/18 20:39:33 mathis Exp S

Done



# Reactive Performance Monitoring

- Current Model
  - User complains about performance to a site
  - Administrator runs 'iperf' or other performance monitoring tools when the complaint comes in.
- Problems with current 'reactive' monitoring
  - Assumes users even know to complain.
  - No way to know how long the problem has occurred.
  - Unlikely to catch intermittent problems

# Regular Testing

- perfSONAR-PS tools eases regular testing
  - Automate running regular throughput and latency tests, and archives the results
  - Allows easy visualization of the history of throughput and latency performance
- pS-Performance Toolkit
  - Bootable Linux ISO with perfSONAR-PS tools
    - bwctl, owamp, NDT, NPAD, PingER, perfSONAR-BUOY (regular tester), Cacti
  - Upcoming release v3.1 in September
    - 'release candidate' versions are being tested

# pS-Performance Toolkit – Test Setup

Throughput tests will be running 2% of the time

## Scheduled Tests

One-Way Delay Test (perfSONAR-BUOY/OWAMP)	<a href="#">Configure</a> <a href="#">Delete</a>
Throughput Test (perfSONAR-BUOY/bwctl)	<a href="#">Configure</a> <a href="#">Delete</a>
PingER Test	<a href="#">Configure</a> <a href="#">Delete</a>

[Add New Throughput Test](#) [Add New Ping Test](#) [Add New One-Way Delay Test](#)

## Test Parameters

Description	One-Way Delay Test (perfSONAR-BUOY/OWAMP)
Packet Rate (packets per second)	10
Packet Size (bytes)	20
<a href="#">Edit Test Parameters</a>	

## Test Members

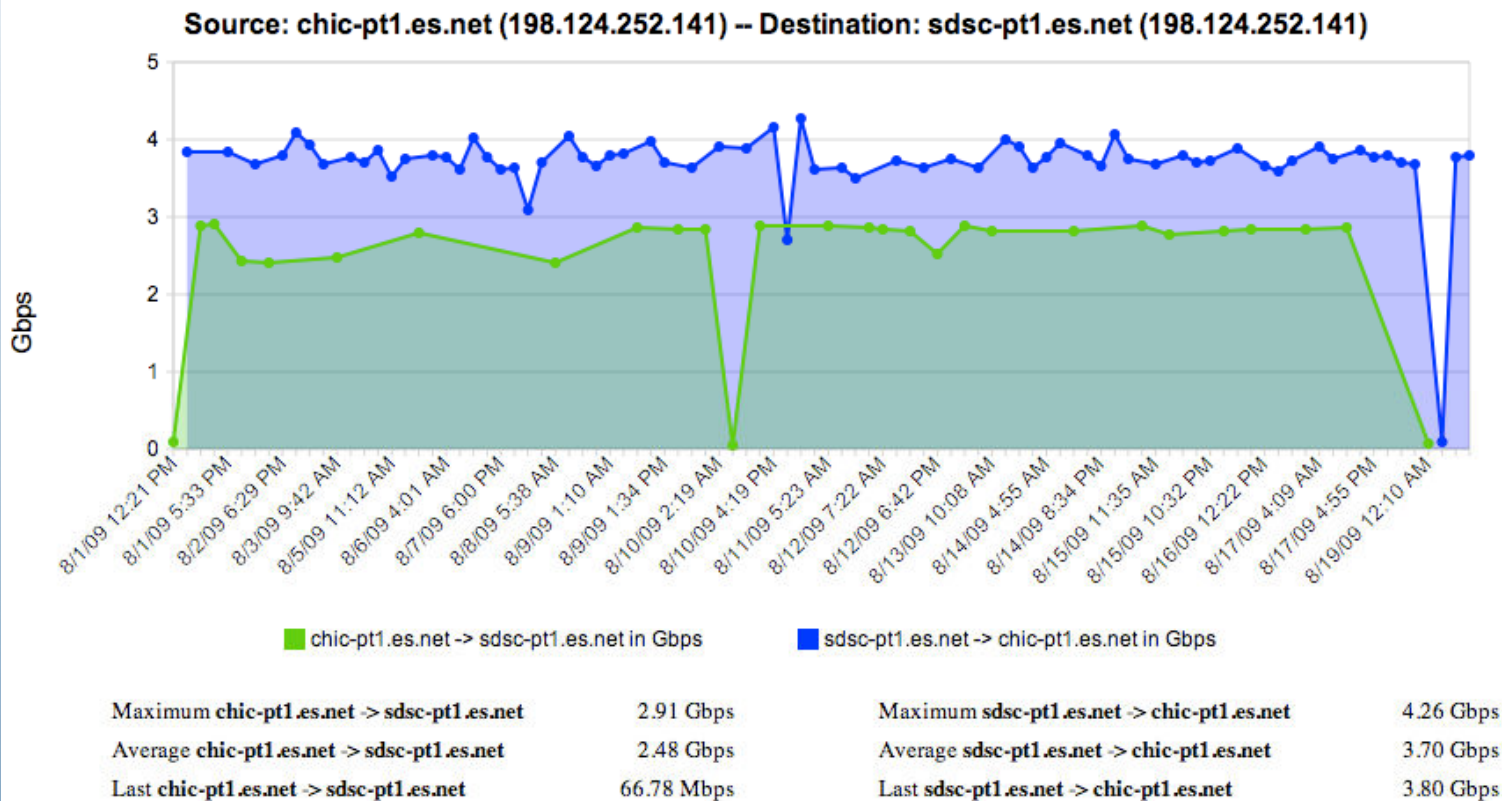
207.75.165.145	Internet2 AA Office	<a href="#">Delete</a>
psum01.aglt2.org	OWAMP Server at University of Michigan in Ann Arbor	<a href="#">Delete</a>
192.52.179.221	Internet2 DC Office	<a href="#">Delete</a>

[Add New Host](#)

## Find Hosts To Test With

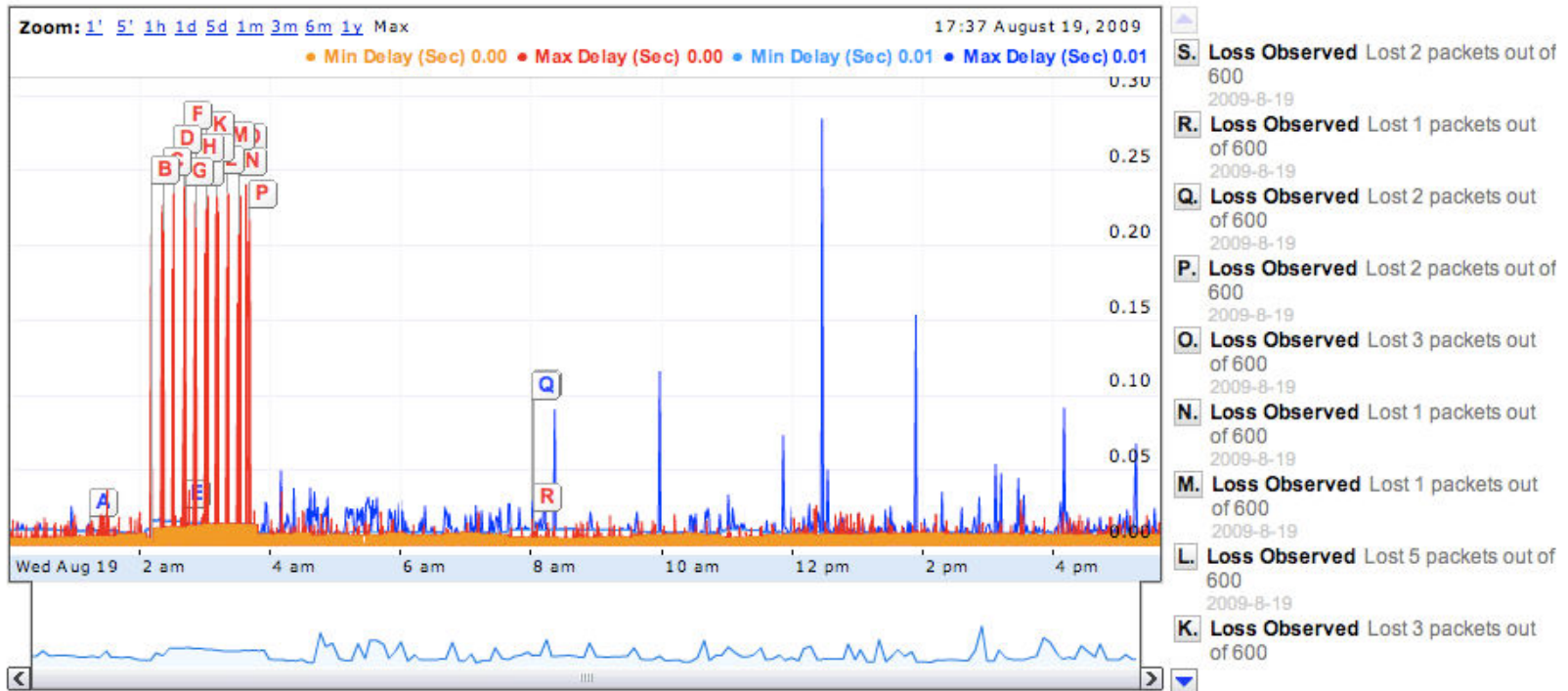
Members Of USATLAS Community As of Wed Aug 19 17:01:25 2009 UTC	
OWAMP Server at Boston University in Bonston, MA, USA	
192.5.207.251	<a href="#">Add To Test</a>
OWAMP Server at OU_OCHEP_SWT2 in Norman, OK, U.S.A.	
129.15.40.232(ps2.ochep.ou.edu)	<a href="#">Add To Test</a>
OWAMP Server at Southern Methodist University Physics Department in Dallas, Texas, USA	

# pS-Performance Toolkit – Throughput



# pS-Performance Toolkit – One-Way Delay

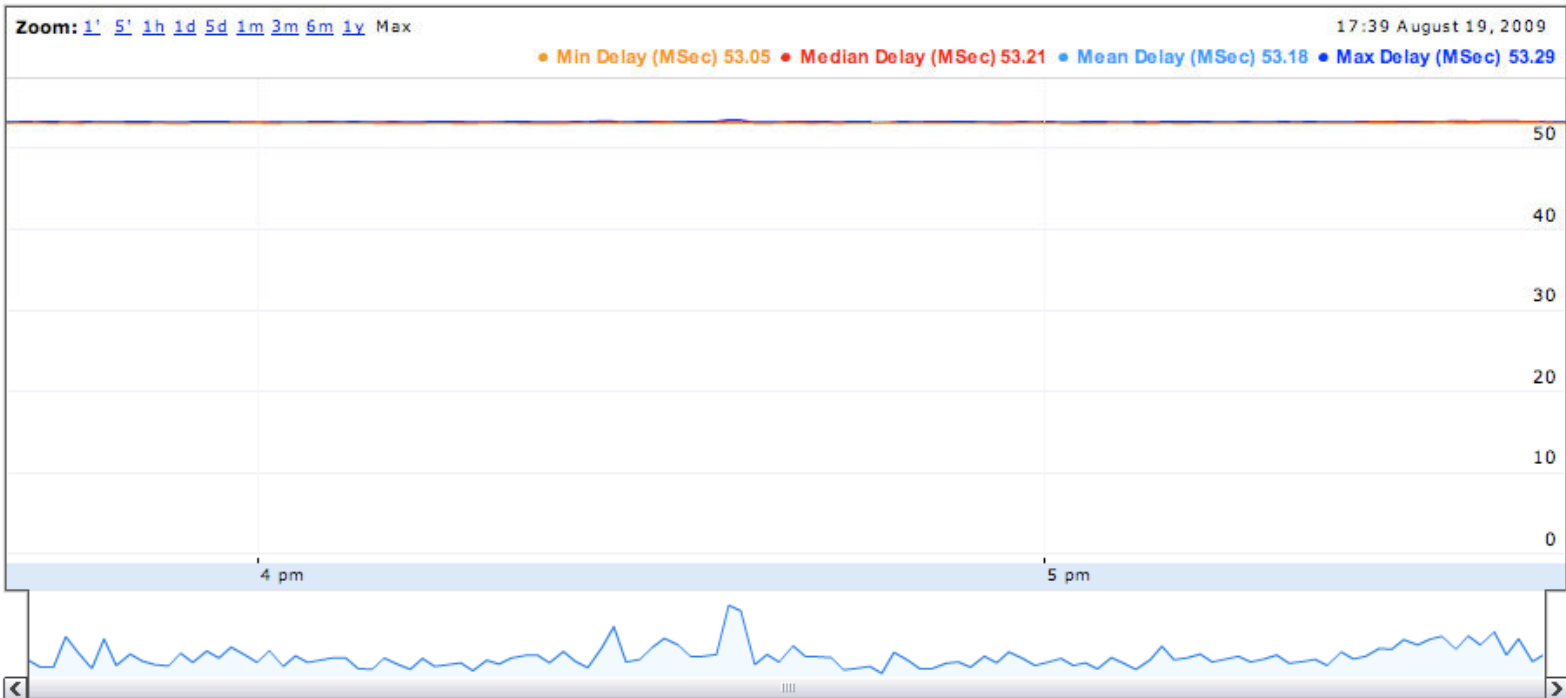
Source: lab246.internet2.edu (207.75.164.246) -- Destination: perfmon-dc.internet2.edu (192.52.179.221)





# pS-Performance Toolkit – Ping Delay

Source: nersc-owamp.es.net (198.129.254.34) -- Destination: star-owamp.es.net (198.124.252.106)



# Testing Recommendations for Tier-2s

- Regular Tests
  - Throughput (bwctl test)
    - 4-hourly 20 second TCP tests to all other Tier-2s and the Tier-1
  - One-Way Jitter/Latency (owamp test)
    - Continuous 10 packet per second tests to all other Tier-2s and the Tier-1
  - Connectivity (ping test using PingER)
    - 5-minutely ping test of 10 packets with a 1 second inter-packet interval

# Testing Recommendations for Tier-2s

- Hardware
  - Two Monitoring Machines
    - One for latency tests and one for throughput tests
    - Throughput tests on the same machine as latency tests will perturb the latency tests
  - Located as closely to resources of interest as possible
- Software
  - pS-Performance Toolkit v3.1
    - New release candidate next week (Aug 24-29)

# Testing Recommendations for Tier-3s

- Still under development
  - Using the Tier-2 deployment experience as input into these testing recommendations.
  - Can use the Tier-2 hardware recommendation
    - <http://code.google.com/p/perfsonar-ops/wiki/Tier2HardwareRecomendations>

# Conclusions

- Tier-2s are deploying the pS-Performance Toolkit to perform regular throughput and latency testing
- This deployment will enable Tier-3s to run on-demand network measurements tests to the Tier-2s
- Eventually, the Tier-3s will be able to perform regular throughput and latency tests with their Tier-2

# Thanks

- Brookhaven National Laboratory (BNL)
- Caltech
- CERN
- ESnet
- Fermilab
- Georgia Institute of Technology
- Indiana University
- Michigan State University
- Mid-Atlantic Crossroads (MAX)
- MCNC
- National Energy Research Scientific Computing Center (NERSC)
- Pacific Northwest National Laboratory (PNNL)
- The Pennsylvania State University
- RNP
- SLAC
- Texas A & M University
- University of Delaware
- University of Michigan
- University of North Carolina
- University of Oklahoma
- Vanderbilt University